Biodiversity Data Journal 7: e36783
doi: 10.3897/BDJ.7.e36783

Research Article

A deafening silence: a lack of data and reproducibility in published bioacoustics research?

Ed Baker‡,§, Sarah Vincent‡

‡ Natural History Museum, London, United Kingdom
§ University of York, York, United Kingdom

Corresponding author: Ed Baker (edwbaker@gmail.com)

Academic editor: Therese Catanach

Received: 04 Jun 2019 | Accepted: 29 Oct 2019 | Published: 30 Oct 2019

Citation: Baker E, Vincent S (2019) A deafening silence: a lack of data and reproducibility in published bioacoustics research? Biodiversity Data Journal 7: e36783. https://doi.org/10.3897/BDJ.7.e36783

Abstract

A study of 100 papers from five journals that make use of bioacoustic recordings shows that only a minority (21%) deposit any of the recordings in a repository, supplementary materials section or a personal website. This lack of deposition hinders re-use of the raw data by other researchers, prevents the reproduction of a project's analyses and confirmation of its findings and impedes progress within the broader bioacoustics community. We make some recommendations for researchers interested in depositing their data.

Keywords

bioacoustics, open data, reproducibility, sound libraries, acoustic vouchers

Introduction

The importance of sharing the datasets used in biological research has been discussed recently by a number of authors, for example, in ecology (e.g. Poisot et al. 2013, Kenall et al. 2014), phylogenetics (e.g. Magee et al. 2014, Stoltzfus et al. 2012) and behaviour (e.g. Caetano and Aisenberg 2014). These authors list several benefits of sharing data,

© Baker E, Vincent S. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

including the opportunity to create future collaborations and clarification of authorship. There can also be ethical considerations, including the use of public funds to generate these datasets. Another significant reason for sharing datasets underpinning research is to ensure that those findings are reproducible, a fact which advocates for open science have discussed well before the recent 'reproducibility crisis' in psychology (Open Science Collaboration 2015).

The datasets used in bioacoustic research vary in scale from a single short recording to continuous recordings at a site over several years. These recordings may be used to identify (e.g. Heller and Baker 2017) or describe (e.g. Dring 1983) species new to science and to facilitate long-duration biological surveys (e.g. Eichinski and Roe 2017). Additionally, recordings may be used to design automated surveys that do not themselves make recordings (e.g. Bennett et al. 2015).

Material and methods

The twenty most recently published articles (as of the end of 2017) covering bioacoustic topics were selected from each of the journals Bioacoustics, ZooKeys, ZooTaxa, Journal of Animal Behaviour and Marine Mammal Science. Primary research articles were identified using the search terms "acoustics" and "bioacoustics" on the journal's website; only articles making use of recorded sounds were selected. The journals chosen cover a subject-specific journal (Bioacoustics), a modern semantically enhanced (Penev et al. 2010) journal (ZooKeys), a taxonomic mega-journal (ZooTaxa), a journal with a broader zoological scope (Journal of Animal Behaviour) and a taxonomically focussed journal (Marine Mammal Science).

Both authors independently examined the papers for evidence that the underlying sound recordings were available in either physical or digital repositories. Two scoring systems were used to categorise the papers:

- The first identifies those papers that define a repository for sound recordings;
- The second differentiates between the type of deposition (repository, personal website, supplementary material) and whether the recordings deposited are the complete set or a representative sample.

Data resources

A CSV file of the papers analysed (including their DOIs) and scored values is available in Suppl. materials 1, 2, full Crossref metadata for the articles is available in Suppl. materials 3, 4.